

A multiple criteria sorting method in a web-based platform: Portuguese public hospital performance assessment

Ricardo Jorge Tomás Oliveira
ricardo.jorge.tomas.oliveira@tecnico.ulisboa.pt

Instituto Superior Técnico, Lisbon, Portugal

November 2019

Abstract

Multiple Criteria Decision Analysis (MCDA) methods are widely used in several fields and disciplines, such as healthcare, and their use can be quite complex. Therefore, a Decision Support System (DSS) can be useful to compute these methods and visualize their results. However, most of the available solutions are desktop based applications with usability issues that not fulfill the requirements of a case study. For this purpose, this work describes the design and implementation of ELECTRE TRI-nC method, a multiple criteria sorting method for handling categorical classification problems taking into account several reference actions to characterize each category, in DECSPACE, a web-based innovative platform for supporting decision aiding processes using one or more MCDA methods in a user-friendly interface. The method implementation was validated by testing it against a numerical example. The demonstration of its use in the platform to hospital performance assessment allows to draw conclusions and recommendations and demonstrates how a DSS can be used to facilitate the process of applying the method in real cases. DECSPACE aggregates the best features of the previous tools, although it stills has margin to be improved in future works, either related to this method or to new ones. Multiple applications of this method in different areas can be carried out by making use of this successfully implementation in DECSPACE.

Keywords: Decision Support System, DecSpace, ELECTRE TRI-nC, Hospital Performance Assessment, Multiple Criteria Decision Analysis.

1. Introduction

MCDA is a sub-discipline of Operations Research that studies the process of using multiple criteria models to help obtain responses to the questions in a decision process. MCDA is not just a collection of formulations and methodologies, but a specific perspective to deal with decision problems [1]. MCDA perspective is closely related to the humans intuition to make decisions. Therefore, the main steps of MCDA can be defined as follows: 1. Defining the decision problem; 2. Selecting and structuring criteria; 3. Measuring performance; 4. Assessing actions/alternatives; 5. Weighting criteria; 6. Aggregating; 7. Dealing with uncertainty; 8. Reporting and examining of findings [2]. This process helps making decisions mainly in terms of choosing, sorting or ranking the alternatives, briefly described as follows:

- *Choice problematic* is focused on the selection of a small number as possible of alternatives in such a way that a single alternative may finally be chosen;

- *Sorting problematic* is focused on an assignment of each action to the most appropriate category among those of a family of predefined categories;
- *Ranking problematic* is focused on a complete or partial order of the actions by comparing actions among each other.

1.1. MCDA in healthcare

Healthcare decisions are complex and involve confronting trade-offs between various, often conflicting, criteria. The iron triangle describes how the three fundamental criteria in healthcare delivery (cost, quality and access) interact: in order to increase one criterion, one or both of the remaining criteria must decrease [3]. Using structured models to improve the quality of decision making involving multiple criteria are useful for this purpose. Some examples of its use in healthcare are: benefit risk analysis, Health Technology Assessment (HTA), resource allocation, portfolio decision analysis, shared patient clinician decision making and prioritizing patients' access to services [4].

Health interventions may be chosen to maximize

general population health and to reduce health inequalities, all with respect to practical and budgetary constraints. This is the type of problem that policy makers need help to solve rationally. This stresses the need for MCDA to support priority setting, which has recently been identified as one of the most important issues in health system research [2].

One of the areas where MCDA is mostly used is HTA, that concerns the way that the costs and benefits of new healthcare technologies are identified and compared. Cost effectiveness is widely used as the criterion, to choose about where to invest on or not. This approach just tells to Decision Maker (DM) about the cost per unit of benefit produced. However, it is not sufficient to make the best decisions: this criterion is just one in many others such as need, appropriateness, quality and ethical or social values [5].

1.2. MCDA-ULAVAL

MCDA-ULAVAL is a free, standalone and open-source desktop application. It only makes available methods from ELECTRE family (specifically ELECTRE II, III, TRI-B, TRI-C, TRI-rC and TRI-nC) and it is the most used tool in the application of methods of this family.

A summary of MCDA-ULAVAL main characteristics (pros and cons) is displayed in Table 1.

Aside from this tool, there are dozens of other DSSs developed with different features and methods available. None of these software tools are such that users without any prior experience of MCDA could use it. An increasing trend is to reduce the gap between researchers and practitioners [6].

The DECISION DECK project was created to collaboratively develop open source software tools implementing MCDA techniques to support complex decision aiding processes. DIVIZ, one of the software initiatives of the project, eases the use of algorithmic resources to build, execute and share complex workflows of MCDA algorithms as ELECTRE TRI family. It is a pedagogical tool for teachers who need to present and compare classical MCDA methods and it may also be used by practitioners who wish to solve decision problems with a given method [7]. However, it also suffers from the flaw of being a desktop standalone application, as MCDA-ULAVAL.

A newly tool in development, DECSPACE (Decision Space), inspired by the workflow design of DIVIZ, is an open source web-based platform optimized to different types of devices using state-of-the-art technology in terms of user interface and back-end that allows a project to be accessed by either non-expert or MCDA expert users, with re-

searching, teaching or even for consulting purposes [8].

In this work, the focus is the implementation of the ELECTRE TRI-nC method in DECSPACE, and its demonstration in the case of Portuguese public hospital performance assessment to draw conclusions and recommendations and to support the application of ELECTRE TRI-nC in DECSPACE for real cases. The Section 3 is dedicated to this platform.

2. ELECTRE TRI-nC method

The ELECTRE TRI-nC method was proposed as a multiple criteria sorting method which takes into account several reference actions to characterize each category [9]. To apply this method, the set of categories to which the actions must be assigned to has to be completely ordered (from the worst to the best, for instance) and defined *a priori*. Each category has to be characterized by several reference actions judged by the DM as representative or informative. This method follows a decision aiding constructive approach, through the interaction between the analyst(s) and the DMs. In the case of each category being only defined by a single reference action, the method is equivalent to ELECTRE TRI-C method [10].

Many applications of these methods were verified in the last years. For instance, the ELECTRE TRI-C method was applied to assisted reproduction [11], for erosion risk assessment [12] and for government performance assessment [13]. The ELECTRE TRI-nC method was used for identifying favourable climates for tourism [14], for supplier classification [15] and for water utilities performance assessment [16].

2.1. Concepts, definitions, and notation

This subsection is dedicated to the main concepts, definitions, and notation concerning the ELECTRE TRI-nC method, adapted from [9] and [10].

Let $A = \{a_1, a_2, \dots, a_i, \dots\}$ denote the *set of actions*. These actions can be known *a priori* or it may appear during the decision aiding process. The main objective of the method is to assign these actions to a *set of categories* $C = \{C_1, C_2, \dots, C_h, \dots, C_q\}$ with $q \geq 2$ (with $q = 1$ there is no sorting problem). The set of categories has to be completely ordered, such that C_1 is the worst category and the C_q is the best one. To evaluate any action, a *set of criteria* $G = \{g_1, g_2, \dots, g_j, \dots, g_n\}$ is defined. $g_j(a_i)$ denotes the performance of the action a_i in the criterion g_j .

In what follows, assume, without loss of generality, that all criteria $g_j \in G$ are to be maximized, which means that the preference increases when the criterion performance increases too. Due to imperfect character of the data as well as the ar-

Table 1: Summary of MCDA-ULAVAL characteristics.

Pros	Cons
✓ Open Source	× Desktop application
✓ Data importation in CSV format	× Not modern user interface
✓ Predefined tables and buttons	× Not adaptable for different types of devices
✓ Data validation	× Just one family of methods available
✓ Scenario analysis	× Not allow multi-user projects

bitrariness that affects the definition of the criteria, two thresholds are associated to g_j : an *indifference threshold*, q_j , and a *preference threshold*, p_j . A *veto threshold*, v_j , can also be associated, such that $v_j \geq p_j \geq q_j \geq 0$. Based on the definition of such thresholds, the following relations between two actions, a and a' , can be derived:

- (i) $|g_j(a) - g_j(a')| \leq q_j$ represents a non-significant advantage of one of the two actions over the other, meaning that a is indifferent to a' according to g_j , denoted $aI_j a'$.
- (ii) $g_j(a) - g_j(a') > p_j$ represents a significant advantage of a over a' , meaning that a is preferred to a' according to g_j , denoted $aP_j a'$.
- (iii) $q_j < g_j(a) - g_j(a') \leq p_j$ represents an ambiguity zone: the advantage of a over a' is enough to not have indifference, but not enough to conclude about a preference in favor of a , meaning that a is *weakly preferred* to a' , denoted $aQ_j a'$.

To quantify the relations between two actions, a and a' , according to a criterion g_j , a *partial concordance index*, denoted $c_j(a, a')$, is defined as follows:

$$c_j(a, a') = \begin{cases} 1 & \text{if } g_j(a) - g_j(a') \geq -q_j, \\ \frac{g_j(a) - g_j(a') + p_j}{p_j - q_j} & \text{if } -p_j \leq g_j(a) - g_j(a') < -q_j, \\ 0 & \text{if } g_j(a) - g_j(a') < -p_j. \end{cases} \quad (1)$$

And a similar definition of *partial discordance index*, denoted $d_j(a, a')$, is defined as follows:

$$d_j(a, a') = \begin{cases} 0 & \text{if } g_j(a) - g_j(a') \geq -p_j, \\ \frac{g_j(a) - g_j(a') + p_j}{p_j - v_j} & \text{if } -v_j \leq g_j(a) - g_j(a') < -p_j, \\ 1 & \text{if } g_j(a) - g_j(a') < -v_j. \end{cases} \quad (2)$$

These partial indices $\in [0, 1]$.

A *criterion weight*, denoted w_j , with $w_j > 0$ is associated to each criterion. Assume, without loss of generality, that $\sum_{j=1}^n w_j = 1$. Thus, the *global concordance index*, denoted $c(a, a')$, is defined as follows:

$$c(a, a') = \sum_{j=1}^n w_j c_j(a, a'). \quad (3)$$

Finally, the *credibility index*, denoted $\sigma(a, a')$, which quantifies the relation between a and a' when taking all the criteria from G into account, is defined as follows:

$$\sigma(a, a') = c(a, a') \prod_{j=1}^n T_j(a, a'), \quad (4)$$

where

$$T_j(a, a') = \begin{cases} \frac{1 - d_j(a, a')}{1 - c(a, a')} & \text{if } d_j(a, a') > c(a, a'), \\ 1 & \text{otherwise.} \end{cases} \quad (5)$$

Let $B = \{B_1, B_2, \dots, B_h, \dots, B_q\}$ denote the *set of subsets of reference actions*. Each *subset of reference actions* to characterize category C_h , is defined as $B_h = \{b_h^r, r = 1, \dots, m_h\}$, such that $m_h \geq 1$. The set $B \cup \{B_0, B_{q+1}\}$ denotes the *set of all reference actions*, such that $B_0 = \{b_0^1\}$ and $B_{q+1} = \{b_{q+1}^1\}$ contain two reference actions defined as follows: $g_j(b_0^1)$ is the worst possible performance on criterion g_j and $g_j(b_{q+1}^1)$ is the best possible performance on criterion g_j for all $g_j \in G$. Using this definition, for any action and criterion, the relation $g_j(b_0^1) < g_j(a) < g_j(b_{q+1}^1)$ is verified.

The comparison of an action a to the characteristic reference actions b_h^r , provides m_h credibility indices $\sigma(a, b_h^r)$ and m_h credibility indices $\sigma(b_h^r, a)$. In order to find a representative credibility index for each action a with respect to each subset of reference actions, B_h , a *categorical index* is defined as follows:

$$\begin{aligned} (i) \quad \sigma(a, B_h) &= \max_{r=1, \dots, m_h} \sigma(a, b_h^r), \\ (ii) \quad \sigma(B_h, a) &= \max_{r=1, \dots, m_h} \sigma(b_h^r, a). \end{aligned} \quad (6)$$

Let λ denote a *credibility level* as the minimum degree of credibility, which is considered necessary by the DM for accept (or not) the outranking statement " a outranks B_h " taking into account all the criteria from G . Normally, $\lambda \in [0.5, 1]$.

2.2. Assignment procedures

This subsection presents the two assignment rules, which must be used conjointly. To apply the two rules, a credibility level is chosen and, to select between two consecutive categories, a *select-*

ing function, denoted $\rho(a, B_h)$, can be defined as follows:

$$\rho(a, B_h) = \min\{\sigma(a, B_h), \sigma(B_h, a)\}. \quad (7)$$

(Descending rule). Decrease h of B_h from $(q+1)$ until the first value, t , such that $\sigma(a, B_t) \geq \lambda$ (C_t is the descending pre-selected category):

- (i) For $t = q$, select C_q as a possible category to assign action a .
- (ii) For $0 < t < q$, if $\rho(a, B_t) > \rho(a, B_{t+1})$, select C_t as a possible category to assign a , otherwise, select C_{t+1} .
- (iii) For $t = 0$, select C_1 as a possible category to assign a .

(Ascending rule). Increase h of B_h from zero until the first value, k , such that $\sigma(B_k, a) \geq \lambda$ (C_k is the ascending pre-selected category):

- (i) For $k = 1$, select C_1 as a possible category to assign action a .
- (ii) For $1 < k < (q + 1)$, if $\rho(a, B_k) > \rho(a, B_{k-1})$, select C_k as a possible category to assign a , otherwise, select C_{k-1} .
- (iii) For $k = (q + 1)$, select C_q as a possible category to assign a .

The assignment procedures lead to select two possible categories to which an action can be assigned to by using the descending and ascending rules conjointly. Therefore, the ELECTRE TRI-nC method provides as a possible assignment of an action:

- (i) One category, when the two selected categories are the same.
- (ii) Two categories, when the two selected categories are consecutive.
- (iii) A range of categories, delimited by the two selected categories when they are non-consecutive.

3. Implementation

3.1. DECSPACE overview

DECSPACE¹ is currently in pre-alpha, meaning that the platform has been in development, and this is an early release (Figure 1).

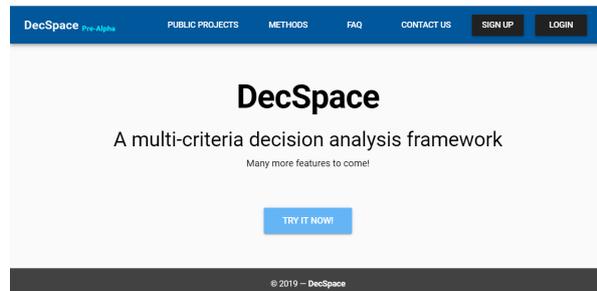


Figure 1: DECSPACE homepage (pre-alpha version).

¹DECSPACE pre-alpha is available at <http://app.decspacedev.sysresearch.org>

DECSPACE consists of a three-tiers architecture:

1. *Client tier*: it implements the user interface and sends user simple HTTP requests to the application tier. After these requests are processed a response is returned to the client.
2. *Application tier*: confines the great complexity of the system, where most of the computational activity is performed. It processes the user requests sent by the client tier, and also carries out the connections to the data tier.
3. *Data tier*: it stores all the information of the application in a database. It receives and replies to any data requests sent by the application tier.

Regarding the technology, a combination of programming languages and libraries is used, through a MEVN stack:

- MongoDB: document database used by the back-end application to store its data as JSON documents.
- Express: framework running on top of Node.js that provides a robust set of features for web and mobile applications.
- Vue.js: front-end web app framework that runs the JavaScript code in the user's browser, allowing the application user interface to be dynamic.
- Node.js: environment that lets implement the application back-end in JavaScript.

Due to the flexible nature of the three-tier architecture, it is possible to migrate the technology of just one of the layers. For instance, the recently front-end migration involved the migration from AngularJS to Vue.js [17].

To start using DECSPACE, there are four types of users with different permissions:

1. *Registered users*: they have an account to create and manage projects with all the features available in a project area.
2. *Anonymous users*: they can test the platform with temporary projects that cannot be saved but can be exported.
3. *Administrators*: they can manage the platform having permission to modify or delete any registered users and projects.
4. *Developers*: they can implement and add new methods to the catalog.

The *Method Catalog* has the available methods created by the developers of the framework to use in projects. Each project has its own information, including the privacy setting, public or private. The *My Projects* area allows the user to create and manage projects. The *Public Projects* area provides access to public projects that were shared by other users. They can be opened and duplicated to other projects. Private projects are just available for the own user.

Any method on the catalog can be used in the

workspace area, by dragging, dropping and properly connecting themselves and data modules in an intuitive and interactive graphical user interface that allows the user to get feedback. This is the most important area, where all the technical work happens. These workflows can be saved, executed, deleted and exported. The user can also select the project version to retrieve a previous version saved.

The input data and preference information can be manually provided in the modules or be imported as a CSV and JSON files. A .zip file can also be imported, containing a workflow that was already used and exported.

3.2. ELECTRE TRI-nC implementation

A MCDA method in DECSPACE can be split in three main components:

1. *Module*: the module is a small, draggable box that can be moved and dropped inside the workspace. A module has a name, a delete button, input slots and an output slot.
2. *Modal*: the modal is a window with a graphical interface specifically designed for each method. It can be accessed by clicking one of the input slots of the module. For most methods, an interactive table should be enough for manual data input, but more complex and dynamic interfaces can be implemented.
3. *Service*: the service represents the algorithm of the MCDA method itself. It receives as input the parameters of that method and when the user executes the workflow it returns an output.

The folders and files that have to be taken into account are the following:

- /client/src/components/Workspace/Methods/: this is where the module and modal files of the method are placed.
- /client/src/services/: this is where the algorithm implementation of the method is placed.
- /client/src/components/Workspace/Workspace.vue: this is the main workspace controller, where few lines of code to link the method are added.
- /client/src/components/Catalog/Catalog.vue: this is where the method and some basic information about it are added to the method catalog.

In order to use ELECTRE TRI-nC in DECSPACE, the user after creating a project, that is displayed in the workspace, must choose the ELECTRE TRI-nC method among the available methods by clicking on the *Add Method* button of the *Method Selection* menu. Immediately, the method module appears in the workspace area. To create the workflow, all required data and preference parameters need to be linked to that module. This can be done

by manually inserting the data in the modal, after clicking on the buttons of the module, or by uploading correctly structured CSV or JSON files, which appear as data modules. These files must have the needed data and table headers (if applied), to be then properly connected to the method module.

The input block contains the main data (actions, criteria and performance table) and the preference parameters (reference actions, criteria weights, thresholds and credibility level), which are contained in the method module.

More specifically, the user can provide the following input information to the method module:

1. *Data*:
 - (a) *Actions*. In this table, the main information about an action is needed:
 - *Name*: it is the name of the action.
 - *Description*: it is the respective action description (not mandatory).
 - (b) *Criteria*. In this table, the following columns appear by default:
 - *Name*: it is the name of the criterion.
 - *Description*: it is the respective criterion description (not mandatory).
 - *Direction*: it corresponds to the criterion preference direction (*Maximize* or *Minimize*).
 - *Scale Type*: it is related to criterion performance levels (*Cardinal* or *Ordinal*).
 - *Min*: if the scale type is *Cardinal*, then a minimum value for the performance levels should be provided.
 - *Max*: if the scale type is *Cardinal*, then a maximum value for the performance levels should be provided.
 - *Num Levels*: If the scale type is *Ordinal*, then the total number of scale levels should be provided.
 - (c) *Performance Table*. The rows correspond to the actions names and the columns corresponds to the criteria names. Performance levels to each criterion must be provided for each action (the platform verifies if they respect the criteria scales characteristics).
2. *Preference Parameters*:
 - (a) *Reference Actions*. For each reference action, the following information is needed:
 - *Category*: the name of the category it belongs to.
 - *Name*: the name of the reference action.
 - *Performance Levels*: performance levels have to be fulfilled in each criterion column.
 - (b) *Criteria Weights*. For each criterion, a value of the criterion weight must be given.
 - (c) *Thresholds*. For each criterion, the prefer-

ence and indifference thresholds must be defined. The veto threshold is not mandatory.

- (d) *Credibility Level*. A value between $[0.5, 1]$ has to be given.

The constructed workflow (Figure 2) can be executed when all connections are properly done and the required data are provided. If the required data has some errors, a message appears to the user. Executing the constructed workflow, the results obtained (method output) appear as a data module in the workspace that contains a table with the classification of the actions into the considered categories. By clicking on the output module box, the corresponding results can be visualized and analyzed by the user. Note that this output module can be used as input module of another method.

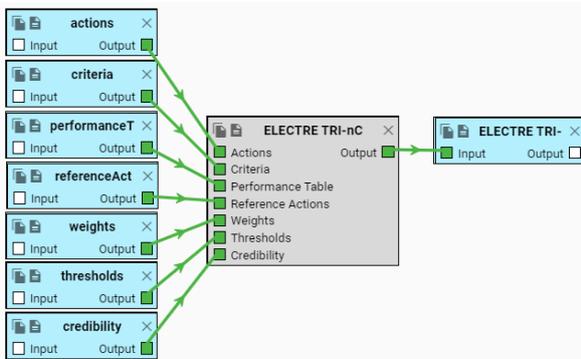


Figure 2: DECSpace workflow of ELECTRE TRI-nC method.

4. Demonstration

4.1. Background

The current Portuguese NHS was established in 1979 in order to provide the appropriate and equitable care to all citizens and is tendentiously free (because co-payments are charged to some customers). This means that the system must be financially sustainable, in particular note that the hospitals consume more than half of the public health expenses. In view of that, several health reforms have been introduced to reduce costs and waste of public funds, while the efficiency and the effectiveness of healthcare providers were expected to improve. These reforms include the creation of Public Enterprise (EPE) hospitals, hospital centers, primary care local health units and Public-Private Partnerships (PPP) [18].

Primary care reform local health units were created in 1999 to enhance greater and better communication between primary care and hospitals, through a vertical integration of different levels of care. At the start of the twenty-first century, the government increased private sector involvement in the building, maintaining and operating of health facilities under PPP. The creation of hospital centers allows better coordination between institutions

providing hospital care in the same geographical area. Both hospital centers and hospitals are EPEs, meaning that hospital boards have some autonomy.

Portugal is progressively becoming more transparent with real time information and an extensive information infrastructure, which plays a central role in monitoring the health system performance.

4.2. Performance assessment in healthcare

Performance in healthcare, as in other services industry, can be defined as as appropriate combination of efficiency and effectiveness. To be efficient a healthcare provider must has a certain level of quality care using the minimum combination of resources. Many times, productivity and efficiency are used with the same purpose. Effectiveness, more specifically, evaluates the outcomes of medical care [19].

A way of assessing performance in healthcare is by measuring the quality of medical care according to the three interrelated categories: outcomes, process and structure [20]. CARE APPROPRIATENESS can be identified as an indicator of outcomes and process quality. Appropriateness of healthcare regards the ability of providing patient-centred care services supported by evidence-based medicine [21].

In other hand, the Portuguese NHS provides universal care, therefore a citizen should has access to a particular service whenever it is necessary or intended. To measure access, some dimensions should be taken into account, such as TIMELINESS OF SERVICES and SERVICES AVAILABILITY. Timeliness of services refers to the capacity of delivering healthcare services whenever required in proper time, and services availability regards the existence of disposable resources to be used when necessary [21].

The system must be financially sustainable, hence the problem of ECONOMIC EFFICIENCY can also be considered as a measurable dimension of performance.

Data Envelopment Analysis (DEA) is a linear programming methodology that is the leading performance evaluation approach commonly used to measure the efficiency of hospitals, by evaluating the relative efficiency of hospitals using multiple inputs to produce multiple outputs. A DEA model was used to assess Portuguese public hospital performance in a previous work [22]. In this section, a MCDA sorting method to assign each hospital to the most appropriate category is presented. For this context, the DM chosen is specialized in health administration.

4.3. Data and sample

All data (hospitals, indicators and performance) are collected from the official benchmarking database, maintained by the Portuguese Central Administration of the Health System at <http://benchmarking.acss.min-saude.pt/>.

This database has a total of 43 institutions in 5 groups using hierarchical clustering and a total of 34 indicators grouped in 6 benchmarking dimensions.

The actions to be assigned to the predefined categories are defined as Portuguese public hospitals. The clustering groups are not considered in this analysis because the object of study are hospitals as individual institutions. The local health units include primary care units, so they are not considered. Hospitals managed as PPP are also not considered because they have a considerable lack of public data. The oncology institutions (IPOs), despite being public hospitals, are very specialized institutions with their specific technology of production [23], so they are not considered, as well. This process leads to a sample composed by a total of 28 institutions (7 hospitals and 21 hospital centers). The Portuguese public hospitals under assessment and the corresponding actions notation, a_i with $i = 1, \dots, 28$, are presented in Table 2.

Table 2: Portuguese public hospitals under assessment.

Action	Hospital
a_1	Centro Hospitalar Barreiro/Montijo, EPE
a_2	Centro Hospitalar de Leiria, EPE
a_3	Centro Hospitalar de Lisboa Ocidental, EPE
a_4	Centro Hospitalar de Setúbal, EPE
a_5	Centro Hospitalar do Baixo Vouga, EPE
a_6	Centro Hospitalar do Médio Ave, EPE
a_7	Centro Hospitalar do Oeste, EPE
a_8	Centro Hospitalar e Universitário de Coimbra, EPE
a_9	Centro Hospitalar Entre Douro e Vouga, EPE
a_{10}	Centro Hospitalar Médio Tejo, EPE
a_{11}	Centro Hospitalar Póvoa de Varzim/Vila do Conde, EPE
a_{12}	Centro Hospitalar Tâmega e Sousa, EPE
a_{13}	Centro Hospitalar Tondela-Viseu, EPE
a_{14}	Centro Hospitalar Trás-os-Montes e Alto Douro, EPE
a_{15}	Centro Hospitalar Universitário Cova da Beira, EPE
a_{16}	Centro Hospitalar Universitário de Lisboa Central, EPE
a_{17}	Centro Hospitalar Universitário de São João, EPE
a_{18}	Centro Hospitalar Universitário do Algarve, EPE
a_{19}	Centro Hospitalar Universitário do Porto, EPE
a_{20}	Centro Hospitalar Universitário Lisboa Norte, EPE
a_{21}	Centro Hospitalar Vila Nova de Gaia/Espinho, EPE
a_{22}	Hospital da Senhora da Oliveira, Guimarães, EPE
a_{23}	Hospital Distrital da Figueira da Foz, EPE
a_{24}	Hospital Distrital de Santarém, EPE
a_{25}	Hospital Espírito Santo de Évora, EPE
a_{26}	Hospital Fernando Fonseca, EPE
a_{27}	Hospital Garcia de Orta, EPE
a_{28}	Hospital Santa Maria Maior, EPE

The indicators in the database are organized in individual benchmark dimensions: ACCESS, PERFORMANCE ASSISTANCE, SAFETY, VOLUME AND USAGE, PRODUCTIVITY, and ECONOMIC-

FINANCIAL. There are, at least, two different procedures to deal with these indicators: 1. Considering individual benchmark dimensions as criteria and indicators as sub criteria (criteria are structured hierarchically [24]); 2. Mixing indicators considering that each one describes a criterion, not considering dimensions, and applying the method once to classify the performance in general (criteria at the same level). A parallel study [25] is using the first methodology. Therefore, the second procedure will be adopted in this work to method demonstration.

In this demonstration, the indicators from SAFETY and VOLUME AND USAGE dimensions, as well as childbirth related indicators, are not considered due to their values being almost null. In the ECONOMIC-FINANCIAL dimension the most representative indicator, OPERATION COSTS, was chosen. This process leads to a sample composed by a total of 9 indicators describing 9 criteria to assess the Portuguese public hospitals as presented in Table 3.

The considered criteria follow the established in Subsection 4.2: TIMELINESS OF MEDICAL APPOINTMENTS and TIMELINESS OF SURGERIES are associated to TIMELINESS OF SERVICES; READINESS OF HIP SURGERIES, WAITING TIME BEFORE SURGERY, OUTPATIENT SURGERIES ADEQUACY, READMISSIONS and DELAY OF CARE are associated to CARE APPROPRIATENESS; BED OCCUPANCY is associated to SERVICES AVAILABILITY and OPERATION COSTS to ECONOMIC EFFICIENCY.

The time interval established to be subject of this analysis was the last year: 2018. Just some exceptions occurred due to the lack of data and in the following cases the data used are from the year 2017: $g_8(a_{28})$, $g_3(a_{19})$, $g_6(a_7)$, $g_9(a_i)$ with $i = 1, \dots, 28$. Note that $g_j(a_i)$ denotes the performance of the action a_i in the criterion g_j .

The performance of 28 hospitals in 9 criteria results in a 28×9 matrix, with 252 entries.

4.4. Preference parameters

This method follows a decision aiding constructive approach, through the interaction between the analyst(s) and the DM(s). This interaction is necessary to define the following preference parameters: categories and their reference actions, criteria weights, thresholds and credibility level.

The DM defined five categories a priori: *Very Weak*, *Weak*, *Neutral*, *Good*, *Very Good* and for each one, one or more (characteristic) reference actions and their performance in the criteria were defined. This information is presented in Table 4.

The DECK CARDS METHOD is used to determine the weights of criteria in the ELECTRE type methods using the revised Simos' procedure [26].

Two different scenarios were taken into account:

Table 3: Criteria, indicators and corresponding direction preferences.

Criterion	Indicator	Direction
g_1 : TIMELINESS OF MEDICAL APPOINTMENTS	Number of non-urgent first medical appointments within adequate time per 100 first medical appointments	Maximization
g_2 : TIMELINESS OF SURGERIES	Number of scheduled surgeries within adequate time per 100 surgeries	Maximization
g_3 : READINESS OF HIP SURGERIES	Number of scheduled hip surgeries performed in the first 48 hours per 100 hip surgeries	Maximization
g_4 : WAITING TIME BEFORE SURGERY	Average number of days after admission waiting before surgery	Minimization
g_5 : OUTPATIENT SURGERIES ADEQUACY	Number of outpatient surgeries per 100 potential outpatient procedures	Maximization
g_6 : READMISSIONS	Number of readmissions in 30 days after discharge per 100 patients	Minimization
g_7 : DELAY OF CARE	Number of inpatients staying more than 30 days per 100 admissions	Minimization
g_8 : BED OCCUPANCY	Absolute difference in inpatient bed annual occupancy rate to a reference value of 85%	Minimization
g_9 : OPERATION COSTS	Operation costs per standard patient	Minimization

Table 4: Categories, reference actions and their performance for each criterion.

Category		g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9
Very Weak	b_1	55	65	30	1.4	70	10	6	12	3500
Weak	$b_{2,1}$	70	70	40	1.0	75	9	5	9	3250
	$b_{2,2}$	65	75	40	1.0	75	9	5	9	3250
Neutral	b_3	80	80	50	0.8	80	8	4	6	3100
Good	$b_{4,1}$	90	85	70	0.5	85	7	3	3	2950
	$b_{4,2}$	85	90	70	0.5	85	7	3	3	2950
Very Good	b_5	95	95	80	0.4	90	6	2	0	2800

- SCENARIO 1: more importance to indicator g_9 , because operating costs are the main concern of the NHS sustainability;
- SCENARIO 2: less importance to indicator g_9 , because in an ethical and politically correct view of the NHS the operating costs are the last issue.

The DECK CARDS METHOD is already implemented in DECSPACE and available in *Method Catalog*. The results obtained for the criteria weights are presented in Table 5.

The DM considered that, to associate an action to a category in a criterion, the difference between the action and the reference action of that category must be less than half of the distance between the reference performance of successive categories. The additional veto threshold was assigned with the objective of increasing the power of certain criteria (the most important ones were considered)

and to mitigate compensatory effects. Therefore, this process leads to the defined thresholds displayed in Table 6.

4.5. Results and discussion

The chosen credibility level was $\lambda = 0.65$. The results obtained from DECSPACE for both considered scenarios are displayed in Table 7.

The robustness of the results is demonstrated by analyzing its stability to the change of the preference parameters, namely criteria weights and credibility level. The two scenarios vary in terms of weights, so it allows to test the robustness with the variation of criteria weights. By varying the criteria weights, it was verified that 25% of the assignment results have changed. The importance of the criterion OPERATION COSTS has a small impact in the assignments. For each scenario, the

Table 5: Criteria weights.

Weights	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9
SCENARIO 1	12.91	12.91	6.12	6.12	10.22	12.92	10.22	10.22	18.36
SCENARIO 2	13.97	13.98	8.81	8.81	11.93	13.98	11.93	11.93	4.66

Table 6: Preference, indifference and veto thresholds for each criterion.

Threshold	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9
Indifference, q_j	5	5	5	0.1	2.5	0.5	0.5	1.5	75
Preference, p_j	10	10	10	0.2	5.0	1.0	1.0	3.0	150
Veto, v_j	30	30	—	—	—	3.0	—	—	—

Table 7: Assignment results for both scenarios.

Hospital	SCENARIO 1	SCENARIO 2
a_1	Weak	Weak
a_2	Neutral	Neutral
a_3	Weak	Neutral
a_4	Neutral	Neutral
a_5	Good	Good
a_6	Neutral	Neutral
a_7	Weak	Weak
a_8	Neutral	Weak
a_9	Neutral	Neutral
a_{10}	Weak	Neutral
a_{11}	Good	Very Good
a_{12}	Good	Neutral
a_{13}	Neutral	Neutral
a_{14}	[Weak,Neutral]	[Weak,Neutral]
a_{15}	Neutral	Neutral
a_{16}	Neutral	Neutral
a_{17}	Neutral	Neutral
a_{18}	Weak	Weak
a_{19}	Neutral	Good
a_{20}	Weak	Weak
a_{21}	Neutral	Neutral
a_{22}	Neutral	Neutral
a_{23}	Neutral	Neutral
a_{24}	Weak	Weak
a_{25}	Neutral	Neutral
a_{26}	Good	Neutral
a_{27}	[Weak,Neutral]	[Weak,Neutral]
a_{28}	Good	Good

results stability by changing the credibility level to $\lambda = 0.60$ and $\lambda = 0.70$ is tested. For $\lambda = 0.60$, it was verified that only 7.14% and 0% of the assignment results have changed for SCENARIO 1 and SCENARIO 2, respectively. For $\lambda = 0.70$, it was verified that only 3.57% and 0% of the assignment results have changed for SCENARIO 1 and SCENARIO 2, respectively. Through this analysis, it can be concluded that the results obtained are robust and conclusions can be drawn from this.

The hospitals with the worst performance are the

following: Centro Hospitalar Barreiro/Montijo, Centro Hospitalar do Oeste, Centro Hospitalar Universitário do Algarve, Centro Hospitalar Universitário Lisboa Norte, Hospital Distrital de Santarém; and the hospitals with the best performances are the following: Centro Hospitalar do Baixo Vouga, Centro Hospitalar Póvoa de Varzim/Vila do Conde, Hospital Santa Maria Maior. The profile of high performing hospitals can be used by other hospitals for benchmarking purposes, allowing them to pursue continuous improvement, as a conceptual framework [27]. In addition, the results can also be used to: improve the quality of healthcare services delivered, support management by monitoring performance indicators and optimize hospitals funding based on their performance.

The results obtained are consistent with a previous study [22] that uses a DEA methodology, where Centro Hospitalar Barreiro/Montijo, Centro Hospitalar Universitário do Algarve and Centro Hospitalar Universitário Lisboa Norte were also classified as inefficient and Centro Hospitalar do Baixo Vouga, Centro Hospitalar Póvoa de Varzim/Vila do Conde as efficient, too. Note that in that study the year under analysis was 2016 and the hospitals: Centro Hospitalar do Oeste and Hospital Santa Maria Maior were not considered due to lack of data. The results obtained are not comparable with the results from parallel study [25] that considers criteria structured in a hierarchical way because of the differences in the amount of criteria.

An important issue that should be addressed in future studies is the possible correlation and interaction between some criteria [28]. Another possible interesting study could include the hospitals managed as PPP, in order to compare their performance with the public hospitals performance. However, the definition of criteria where the PPP have public data can be a difficult issue.

5. Conclusion and future work

The main focus of this thesis was the implementation of the ELECTRE TRI-nC method in DECSPACE

and its demonstration to support its application in real cases.

By comparing ELECTRE TRI-nC implementation in DECSPACE versus MCDA-ULAVAL, the first one solves all the cons that was presented to MCDA-ULAVAL. However, the following pros: data validation and scenario analysis are not yet in their full potential in DECSPACE. Therefore, for the future work of ELECTRE TRI-nC in DECSPACE the following hints are presented:

- The data validation of inputs should be improved in order to allow the workflow to be successfully executed only when everything is correctly provided.
- The robustness analysis should be facilitated to allow the user to easily change parameters and see the changes originated.
- The indifference, preference, and veto thresholds have been presented as constants, however, in practice, these thresholds can vary according to the performances $g_j(a)$ [10]. This should be further studied, to understand if the variation should be just linear, and then be implemented in DECSPACE with a user-friendly interface.
- The DECK CARDS METHOD, because is used to determine the criteria weights for ELECTRE methods, should has its output consistent with the ELECTRE TRI-nC implementation to allow it to be the input of criteria weights, by connecting it to ELECTRE TRI-nC module.
- A button to export DECSPACE tables to documents should be developed.
- More ways to visualize the output data, besides the generic data table, should be available to the user. This should be done using D3.js, a JavaScript library to manipulate documents based on data [17].

More MCDA methods and theses should be developed with this platform to continue the work performed and to increase the DECSPACE method catalog. Afterwards, multiple applications of this method in different areas can be applied in the future, taking advantage of its implementation in DECSPACE, thus continuing the research performed in the last years.

Acknowledgements

This document was written and made publically available as an institutional academic requirement and as a part of the evaluation of the MSc thesis in Biomedical Engineering of the author at Instituto Superior Técnico. The work described herein was performed at the Instituto de Engenharia de Sistemas e Computadores Investigação e Desenvolvimento (INESC-ID), Civil Engineering Research and Innovation for Sustainability (CERIS) and Centro de Estudos de Gestão (CEG-IST) of In-

stituto Superior Técnico (Lisbon, Portugal), during the period February-September 2019. This work was part of the hSNS FCT – Research Project (PTDC/EGE-OGE/30546/2017).

References

- [1] J. Figueira, S. Greco, and M. Ehrgott, Eds., *Multiple Criteria Decision Analysis, State of the Art Surveys*. Springer's International Series, 2005.
- [2] P. Thokala, R. Baltussen, T. Lönngren, M. Boysen, N. Devlin, S. Peacock, J. Watkins, K. Marsh, Z. Kaló, M. Ijzerman, and F. Mussen, "Multiple Criteria Decision Analysis for Health Care Decision Making—Emerging Good Practices: Report 2 of the ISPOR MCDA Emerging Good Practices Task Force," *Value in Health*, vol. 19, no. 2, pp. 125–137, 2016.
- [3] D. N. Van Der Goes, N. Edwardson, V. Rayamajhee, C. Hollis, and D. Hunter, "An iron triangle roi model for health care," *ClinicoEconomics and Outcomes Research*, vol. 11, pp. 335–348, 2019.
- [4] P. Thokala, N. Devlin, K. Marsh, R. Baltussen, M. Boysen, Z. Kalo, T. Longrenn, F. Mussen, S. Peacock, J. Watkins, and M. Ijzerman, "Multiple criteria decision analysis for health care decision making - An introduction: Report 1 of the ISPOR MCDA Emerging Good Practices Task Force," *Value in Health*, vol. 19, no. 1, pp. 1–13, 2016.
- [5] N. J. Devlin and J. O. N. Sussex, *Incorporating multiple criteria in HTA*. Office of Health Economics, 2011.
- [6] J. Mustajoki and M. Marttunen, "Comparison of multi-criteria decision analytical software for supporting environmental planning processes," *Environmental Modelling and Software*, vol. 93, pp. 78–91, 2017.
- [7] P. Meyer and S. Bigaret, "Diviz: A software for modeling, processing and sharing algorithmic workflows in MCDA," *Intelligent Decision Technologies*, vol. 6, no. 4, pp. 283–296, 2012.
- [8] A. D. Barbosa, "DecSpace : A Multi-Criteria Decision Analysis Framework," *MSc Dissertation, Instituto Superior Técnico, Universidade de Lisboa*, 2017.
- [9] J. Almeida-Dias, J. R. Figueira, and B. Roy, "A multiple criteria sorting method where each category is characterized by several reference

- actions: The Electre Tri-nC method,” *European Journal of Operational Research*, vol. 217, no. 3, pp. 567–579, 2012.
- [10] —, “Electre Tri-C: A multiple criteria sorting method based on characteristic reference actions,” *European Journal of Operational Research*, vol. 204, no. 3, pp. 565–580, 2010.
- [11] J. R. Figueira, J. Almeida-Dias, S. Matias, B. Roy, M. J. Carvalho, and C. E. Plancha, “Electre Tri-C, a multiple criteria decision aiding sorting model applied to assisted reproduction,” *International Journal of Medical Informatics*, vol. 80, no. 4, pp. 262–273, 2011.
- [12] F. Macary, J. A. Dias, J. R. Figueira, and B. Roy, “A Multiple Criteria Decision Analysis Model Based on ELECTRE TRI-C for Erosion Risk Assessment in Agricultural Areas,” *Environmental Modeling and Assessment*, vol. 19, no. 3, pp. 221–242, 2014.
- [13] A. S. Costa, J. Rui Figueira, C. R. Vieira, and I. V. Vieira, “An application of the ELECTRE TRI-C method to characterize government performance in OECD countries,” *International Transactions in Operational Research*, vol. 26, pp. 1935–1955, 2017.
- [14] D. Mailly, I. Abi-Zeid, and S. Pepin, “A Multi-Criteria Classification Approach for Identifying Favourable Climates for Tourism,” *Journal of Multi-Criteria Decision Analysis*, vol. 21, no. 1, pp. 65–75, 2014.
- [15] A. S. Costa, K. Govindan, and J. R. Figueira, “Supplier classification in emerging economies using the ELECTRE TRI-nC method: A case study considering sustainability aspects,” *Journal of Cleaner Production*, vol. 201, pp. 925–947, 2018.
- [16] F. S. Pinto, A. S. Costa, J. R. Figueira, and R. C. Marques, “The quality of service: An overall performance assessment for water utilities,” *Omega*, vol. 69, pp. 115–125, 2017.
- [17] R. Rodrigues, “Front-end Migration and User Experience Improvements in the DecSpace MCDA Framework,” *MSc Dissertation, Instituto Superior Técnico, Universidade de Lisboa*, 2019.
- [18] J. D. A. Simões, G. F. Augusto, and C. Hernández-Quevedo, *Portugal: Health system review*, vol. 19, no. 2, 2017.
- [19] Y. A. Ozcan, *Health Care Benchmarking and Performance Evaluation*, ser. International Series in Operations Research & Management Science. Springer, vol. 120, 2014.
- [20] A. Donabedian, “Evaluating the quality of medical care,” *The Milbank Quarterly*, vol. 83, no. 4, pp. 691–729, 2005.
- [21] D. C. Ferreira and R. C. Marques, “Do quality and access to hospital services impact on their technical efficiency?” *Omega*, vol. 86, pp. 218–236, 2018.
- [22] M. A. Pereira, “Preference information incorporation for decision-making in a DEA model using the Choquet integral,” *MSc Dissertation, Instituto Superior Técnico, Universidade de Lisboa*, 2018.
- [23] D. C. Ferreira, R. C. Marques, and A. M. Nunes, “Economies of scope in the health sector: The case of Portuguese hospitals,” *European Journal of Operational Research*, vol. 266, no. 2, pp. 716–735, 2018.
- [24] S. Corrente, S. Greco, and R. Słowiński, “Multiple Criteria Hierarchy Process for ELECTRE Tri methods,” *European Journal of Operational Research*, vol. 252, no. 1, pp. 191–203, 2016.
- [25] A. Rocha, “Quality assessment of the Portuguese public hospitals: A multicriteria approach,” *MSc Dissertation, Instituto Superior Técnico, Universidade de Lisboa*, 2019.
- [26] J. Figueira and B. Roy, “Determining the weights of criteria in the ELECTRE type methods with a revised Simos’ procedure,” *European Journal of Operational Research*, vol. 139, no. 2, pp. 317–326, 2002.
- [27] M. Augusto, J. Lisboa, M. Yasin, and J. Figueira, “Benchmarking in a multiple criteria performance context: An application and a conceptual framework,” *European Journal of Operational Research*, vol. 184, pp. 244–254, 2008.
- [28] J. R. Figueira, S. Greco, and B. Roy, “ELECTRE methods with interaction between criteria: An extension of the concordance index,” *European Journal of Operational Research*, vol. 199, no. 2, pp. 478–495, 2009.